DOCUMENT RESUME

ED 259 023 TM 850 393

AUTHOR Gray, Wayne D.
TITLE Implementation Monitoring: A Role for Evaluators in Helping New Programs Succeed.
PUB DATE Oct 84
NOTE 36p.; Paper presented at the Annual Meeting of the Evaluation Research Society (San Francisco, CA, October 10-13, 1984).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Data Interpretation; *Evaluation Methods; Evaluators; Feedback; *Innovation; Models; Program Development; *Program Evaluation; *Program Implementation
IDENTIFIERS Evaluation Problems; *Implementation Analysis; *Monitoring

ABSTRACT

This paper presents a framework for monitoring implementation which defines a process for implementing new programs into organizations. The process requires a team of monitors who examine the adequacy of implementation plans and look at the effect of plan execution upon the organization, individual, and new program. Immediate feedback is provided to adjust the implementation effort to better help the new program succeed. The framework defines questions to ask during the period when implementation is occurring. It shows why these questions are important, how they interrelate, and how they can be used to increase the implementation of innovations as well as to further the study of implementation processes. An optimal use of the framework would be to improve the use of new programs by implementation monitoring. The framework may be used to raise the awareness in organizations, and among evaluators, of the issues involved in implementing new programs. (DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED259023

BEST COPY AVAILABLE

Implementation Monitoring:

A Role for Evaluators in Helping New Programs Succeed

Wayne D. Gray

US Army Research Institute
5001 Eisenhower Ave
Alexandria, Virginia
(202) 274-5779

Presented as part of the panel titled, Implementation: A Key to High Organizational Performance, at the annual meeting of the Evaluation Research Society, San Francisco, CA, October, 1984.

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ The document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Gray, Wayne D.

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

RUNNING HEAD: IMPLEMENTATION MONITORING

Implementation Monitoring:
A Role for Evaluators in Helping New Programs Succeed[1]

Many failures of programs to live up to expectations can be traced to a failure in implementation. Many implementation failures can in turn be traced to the perception of implementation as an event rather than as a critical process in the life cycle of a new program. Viewing implementation as an event leads users to spend time and money on research, development, or purchase of new programs, but nothing on implementation. leads evaluators to attempt summative evaluations of programs which are the midst of the radical change entailed by implementation. In contrast, viewing implementation as a process leads to a consideration of the changes required in both the organization (for example, resources or policies) and individuals (for example, education and acceptance) when a new program is adopted.

There are many program failures which, after the fact, have been traced to a failure of implementation. Severe problems exist even in organizations such as the US Army where outsiders falsely perceive implementation as being easily mandated. For example, in discussing the failure of REALTRAIN, a program for tactical team training, Scott (1983) writes that after delivery of instructor training and equipment,

> REALTRAIN was plunged into a highly complex training environment rife with competing demands for time, personnel and resources. Company and higher level commanders tended to indicate that they had considerable difficulty in meeting the REALTRAIN support

> requirements, especially the requirements for exercise controllers ... additional equipment was required. Although many commanders did not see the additional equipment requirements per se as a major problem, the time and effort required to request, obtain, issue, install, organize, and account for the equipment was often seen as a more serious deterrent to the routine use of REALTRAIN. (p.13-14)

REALTRAIN's effectiveness was never an issue in its implementation. Studies by the Army Research Institute (see Scott, 1983) leave no doubt that when used appropriately REALTRAIN was an astonishingly effective program for tactical team training. It died because it did not "fit" the existing training environment. If the implementation process had been more extensively planned and monitored the disconnect between available resources and REALTRAIN's support requirements would have become an important issue early on. Alternatives did exist. For example, the program could have been changed to minimize support requirements, additional resources could have been provided to local commanders, and streamlined procedures to "request, obtain. . . and account for the equipment" could have been instituted.

The perspective adopted here is that implementation is a process in which evaluators can play an important role: implementation monitoring. The monitor examines the adequacy of implementation plans and looks at the effect of their execution upon the organization, individual, and new program. Immediate feedback is provided to adjust the implementation effort to better help the new program succeed. This

is a more activist role than most evaluators assume. In this paper I describe a framework on which implementation monitoring can be based. As secondary applications, the framework can be used to organize the empirical data concerning implementation processes and to provide a structure for evaluating theories of implementation.

What Is (Not) Discussed

Scheirer and Rezmovic's review (1983) of 74 studies of implementation found a general lack of explicit conceptualization of key terms. To aid communication I will follow Scheirer's exhortations and state here the range and limits of the proposed framework.

First, the framework does not consider adoption processes. Adoption refers to the processes surrounding an organization's decision to try an innovation. Second, the framework does consider programmed implementation (Berman, 1980) . Programmed implementation refers to those changes in the user environment (at both the organizational and individual level) required for the innovation to be used routinely. Third, adaptive implementation (Berman, 1980) is considered. By contrast to programmed implementation, adaptive implementation refers to changes in the innovation required for it to better "fit" the user environment. Finally, the framework considers how much and what parts of the innovation have achieved routine use, that is, degree-of-implementation.

The Framework

An overview of the framework for implementation monitoring is presented as Figure 1. There are three parts to the framework.

Insert Figure 1 about here

Background knowledges are considered as pre-requisite to a monitoring effort. If the monitoring team does not have expertise in the areas listed, then they must have ready access to such expertise.

The rational analyses result in a separation of the program into (more or less) independent components, a rank ordering of these components, an analysis-of-fit between the program and its intended environment, and an analysis of the strategies and tactics needed to overcome potential "misfits". The term "rational" in rational analyses is meant to contrast with "empirical" in empirical evaluations. The analyses are conducted by a team that possesses the relevant background knowledges. Active data collection or statistical analysis is minimal.

The empirical evaluations examine three issues: how well implementation plans (strategies and tactics) are working, how much of the program is currently implemented, and how effective the program is at its current level of implementation. Data is collected as it becomes available. Data analysis is ongoing, but while statistical analyses may be performed, typically preliminary conclusions are made and fed back into the system before enough data exists for statistical techniques to produce reliable results.

The distinction between the three parts of the framework may become clearer by analogy to laboratory experimentation. In this analogy the background knowledge is the experimenter's knowledge of the problem area, previous research, tools available (computers, memory drums, and so on), subject population, and level of funding. The rational analyses correspond to the process of deciding upon an experimental design that will test the hypothesis and that can be conducted with the tools and subjects available. Finally, the empirical evaluations are analogous to data collection and statistical analysis.

Also shown in Figure 1 are the major subparts for each of the three parts and some of their interconnections. To avoid clutter, some of the interconnections and all of the feedback loops have been omitted; however, all parts, subparts, interconnections, and feedback loops are discussed below.

Background Knowledge

As shown, both practical and theoretical knowledge is required to conduct implementation monitoring. For the innovation, practical knowledge of both the procedures involved in using it and the resources needed to support it is required. Practical knowledge of the implementing organization is also required. One or more in-house informants are needed who possess a detailed knowledge of current procedures, resource capabilities and limitations, staff morale and education, key players within the organization, and general organizational climate. Implementation monitoring, therefore, must be a

largely in-house affair (or the monitoring team must work closely with in-house personnel).

In accord with Chen & Rossi's (1983) call for a theory-driven approach to evaluation, two types of theoretical knowledge are required by the framework. The first is an implementation theory. An adequate theory of implementation should provide a comprehensive analytic framework for considering potential implementation problems (for both adaptive and programmed implementation). It should recognize that not all problems have solutions, but provide a clear and prescriptive mapping to solution strategies for those that do. At present, no one theory completely fits these criteria but the ones proposed by Roberts-Gray (Roberts-Gray & T.Gray, 1983; T.Gray, 1981; Roberts-Gray, 1983) and Scheirer (1981; Scheirer & Rezmoric, 1983) are a good beginning.

Second, a theory that covers the domain of the new program (for example, classroom instruction, criminal justice reform, adult literacy, and so on) must be found. Such a theory provides a somewhat objective way of deciding what features of the new program are most important to its effectiveness and what features are merely nice to have. Any theory selected must be appropriate to the innovation and user environment. (This may require using different theories to characterize different innovations even within the same institution and subject domain. For example, for Army instruction, one theory may be appropriate for characterizing tactical field training, another for "schoolhouse" training, a third for "simulator" training, a fourth for CAI, and yet another for the Army's "educational" programs.) Another consideration

is that the theory used should have face validity for the clients and sponsors of the monitoring effort. There are two reasons for this. First, members of the user organization will probably be working with the monitoring team as experts on the organizational and political realities of their organization. Hence the theory-based analysis should divide the programs into components which these experts can readily appreciate. Second, a theory which has face validity makes communication between the monitors and sponsors and users of the research easier. Because less time is required to explain the researcher's assumptions, more time can be spent communicating findings and recommendations.

An example of this is a framework which combines Army guidance on how to conduct field training with the model of teacher behavior proposed by Far West Labs (see for example, Fisher, et al., 1981). (See Table 1.) The headings of preparation, presentation, and practice are taken verbatim from Army guidance on field training. The sub-headings under presentation, practice, diagnosis and prescription are variations

---

Insert Table 1 about here

---

on the model of teacher behavior. Finally, the lower levels translate "guidance" from the teacher behavior model into terms appropriate to tactical field training. Such a framework provides a useful way to analyze the procedures involved in tactical field training programs.

In sum, four types of background knowledge (in addition to

knowledge of evaluation theories and procedures) are required to do implementation monitoring: practical knowledge of the innovation, practical knowledge of the implementing organization, knowledge of domain-relevant theories, and knowledge of implementation theories. This knowledge requirement necessitates a talented team with cross-disciplinary expertise that is part of, or works closely with, the implementing organization.

Rational Analyses

Rational analyses include three activities (see Figure 1): a ranking of program components, an analysis-of-fit, and a characterization of the implementation effort in terms of strategies and tactics. These analyses provides the basis, or design, for the three empirical evaluations. In addition, they provide the basis for a rational (as opposed to empirical) evaluation of both the innovation and implementation planning.

Ranking of program components.

> It is not usually clear whether the recorded failures of programs are due to the fact that the programs were built on poor conceptual foundations . . . or because treatments were set at such low dosage levels that they could not conceivably affect any outcomes . . . or because programs were poorly implemented. Note that the emphasis in the above statements is on deficiencies in the theoretical underpinnings of the treatment or of the treatment delivery systems. (Chen & Rossi, p. 284)

Ranking the components of the program provides a basis for deciding

which components are most important to the program's effectiveness. The goal is to separate the "must have" components from the "nice to have" ones. When resources are limited, this ranking allows implementation efforts to focus on the most important aspects of the new program.

Ranking is a two step process. First, the new program must be separated into (more or less) independent components. In the monitoring framework this separation is based upon background knowledge of the new program with the domain theory providing the organizing framework. For cases where this background knowledge is insufficient, then a more formal procedure such as that discussed by Emshoff et al (1984) is required. An example of a theory-based component analysis is provided in Table 2. The example uses the framework provided in Table 1 (and discussed above) to characterize feedback aspects of an Army program for tactical field training. (More information on this training program can be found in W. Gray, 1983).

Insert Table 2 about here

Second, each component should be ranked as, at a minimum, "must have" or "nice to have." (Generally finer rankings covey more information; however, for most new programs fairly gross rankings should suffice.) This separation into and ranking of components provides input to both the analysis-of-fit and the evaluation of degree-of-implementation. Note that the ranking of program components does not consider the resources each component may require. Resources become an

issue in the analysis-of-fit.

Analysis-of-Fit

The purpose of an Analysis-of-Fit is to identify areas where the routine use of the innovation will require changes in the user's operational environment, that is, places in which the innovation currently "misfits" the organization. The first step in such an analysis is to cross the ranked program components with (relevant parts of) an implementation theory to form a matrix of potential implementation problems. After the matrix is generated, expert opinion (for both the organization and resource needs of the innovation) is required to decide whether any given cell of the matrix "makes sense".

As an example, consider several items from a matrix that was generated by crossing the theory-based component analyses provided in Table 2 with some of Scheirer's (1981) implementation hypotheses. Her hypothesis that it is important for "central administrators (to) strongly and actively support the program" (p.69) does not seem to apply at this level of analysis. Presumably, central administrators support or do not support the program as a whole, not those subcomponents involved in providing feedback to students. On the other hand, the hypothesis that supervisors need to "receive adequate training to understand the philosophy and behaviors specific to the innovation" (p. 70), makes sense to apply. Again, with reference to Table 2, the issue of malfunctioning equipment requiring repair or replacement on-the-spot is different than the prevailing attitudes towards tactical training. Likewise, having the controllers NOT provide feedback on tactical

proficiency during the exercise involves a different philosophy concerning the value of realism and the place for feedback than had previously been accepted. Hence, both these cells are meaningful and should be kept.

As shown by the above examples, making judgments concerning the meaningfulness of any given cell requires expert familiarity with the implementating organization and the resource requirements of the innovation. Also, these two expertises are required to assess whether the meaningful cells represent a match, important misfit, or trivial misfit of the innovation to the organization. If information is lacking concerning the importance of the misfit then a separate data collection effort may be required. Those cells in which the mismatch is deemed important and for which the components are ranked as "must have" represent potential implementation problems (PIPs) that must be solved to implement the new program.

Characterization of the Implementation Effort

The implementation effort can be characterized in terms of strategies and tactics. A strategy is what must be done to overcome a PIP, whereas a tactic is one way of accomplishing the strategy. For example, if computer-based instruction is being implemented and one PIP is instructors' fear of computers, then one strategy to solve this PIP would be to hire instructors who know and love computers. Another strategy would be to educate the current instructors on the advantages of CBI. If the later strategy is choosen, then one tactic might be to enroll all instructors in CBI courses. An alternative tactic might be

to provide instructors with basic reference materials and actively involve them in the development of new courseware. Hence, a strategy is "what to do" while a tactic is "how to do it."

The monitor team may be asked to either develop implementation plans or to clarify (evaluate) existing ones. In either case the activities are about the same (see Figure 2); however, the order of the steps will vary. In developing plans the monitors can follow the steps as listed in Figure 2. In contrast, clarifying existing plans may require a top-down, bottom-up analysis. Starting at the bottom (of Figure 2) with an analysis of the implementation efforts (tactics) and at the top with PIPs the monitor will have to determine what problems the tactics may potentially resolve. Having an implementation theory which maps problems onto solution strategies is helpful.

Insert Figure 2 about here

For example, one PIP in implementing an Army tactical training program might be the redefinition of the controller's role in an exercise as an observer and data collector as opposed to providing on-the-spot feedback and advice concerning tactical performance. As part of the implementation the following efforts are planned: (1) the value of exercise realism is stressed in train-the-trainer programs; (2) program manuals emphasize the importance of the controller's remaining "tactical" during the exercise and collecting information to use in the After Action Review; (3) local training regulations are revised to

eliminate the requirement that controllers wear white hats and belts when conducting an exercise; (4) special manuals are prepared and targetted at Battalion and higher commanders, explaining the philosophy of the new program. All four efforts are examples of tactics taken to resolve the PIP. Tactics 1, 2, and 4 are examples of an education strategy with the goal of educating the controllers and their commanders on the new way of doing things. Tactic 3 is a power strategy (Roberts-Gray & T.Gray, 1983) with the goal of eliminating regulations inconsistent with the new program.

The clarification of what is being done and why has two functions. First, clarification is pre-requisite to evaluating strategies and tactics, and this evaluation provides a reading of the success of the implementation effort in resolving PIPs. For example, to resolve the PIP of role change for controllers, implementation theory (Roberts-Gray & T.Gray, 1983) suggest two strategies: the use of power and education. To determine if the use of power was successful, the strategy of power must be tied to a particular tactic--change of certain regulations. Likewise, the strategy of education must be tied to certain tactics-- train-the-trainer programs, manuals, and officer education. As discussed later, the evaluation of strategies, tactics, and PIPs are interrelated and the relationship of a given tactic to strategy to PIP must be determined for the evaluation to provide useful results.

Second, a rational analysis of the implementation effort will more than likely pinpoint PIPs which are not being addressed and tactics that

serve no purpose, that is, that would not resolve any PIP. Such a clarification represents a rational evaluation of the implementation effort and should be fed back to the implementing organization as soon as possible.

Empirical Evaluations

The empirical evaluations are the heart of implementation monitoring. As shown in Figure 1, the three evaluations are interrelated with evaluation of degree-of-implementation providing inputs into the trial, and strategies and tactics evaluations. Not shown are the feedback loops from these last two evaluations to various issues in rational analyses and background knowledge.

Not discussed here are details concerning how to conduct the various evaluations. The tools to do such evaluations already exist in the evaluator's toolbox. Nothing unusual or esoteric is envisioned. What is discussed are the issues involved in each of the three evaluations, why these issues are important, and how they relate to each other and to issues discussed above. Put another way, the implementation monitoring framework assists the evaluators in deciding what questions to ask and why. How these questions are asked is left to the discretion and inclination of the evaluator.

Evaluating Degree-of-Implementation

Typically, assessing degree-of-implementation is viewed as important to interpreting summative evaluation data (for example, Fullen & Pomfret, 1977; Leinhart, 1980). Hall & Loucks (1977) offer guidelines concerning how such an evaluation might be conducted.

The emphasis of the current framework is different. First, the ranking of program components determines the relative importance of various aspects of the new program (discussed above). Second, the evaluation is started during the implementation process, and long before any complete implementation is expected. Once started the evaluation is continually updated until all implementation efforts have stopped. Third, measurement of degree-of-implementation falls into two different categories, fidelity and sufficiency (see Figure 3). For fidelity the task is to assess what aspects of the innovation have been implemented, what have been dropped, and what have been altered. Sufficiency focuses on those parts which were altered to assess whether the alterations fulfill the function of the omitted part and therefore will support the goals of the innovation.

Fidelity Issues. Measuring fidelity consists of comparing the user's procedures against the developer's ideal.

Insert Figure 3 about here

Components which were dropped provide input to the strategy and tactic evaluations. For these components it is important to determine whether their implementation requires tactics to be revised or new strategic goals set (see Figure 3). In contrast, components which were implemented-as-intended provide input to the trial evaluation. As discussed below, knowledge of which components have been implemented is used to weigh the significance of the current effectiveness of the

program-as-a-whole. Last, if fidelity measures determine that a component has been altered, then the sufficiency of the alteration is assessed.

Sufficiency Issues. In measuring sufficiency, the domain theory that was used to rank the program components is used to judge the alteration. (Note that what I am calling sufficiency issues have a relationship in instruction to what Leinhart (1980) has called Domain-of-Instruction.) The important point is that for many programs, alterations are a fact of life (Berman, 1978); however, alterations per se should not be assumed to be innately bad or good. An alteration is not bad just because it differs from the developer's ideal. It is not good just because it represents an adaptation of the innovation to the local way of doing things. Rather all alterations must be measured by the more objective standard of sufficiency: based upon the domain theory, does the alteration fulfill the function of the omitted part? Will it support the goals of the innovation?

For example, many training programs include procedures to provide feedback to the trainees. However, if the exact procedures specified by the program are not followed, feedback may still be provided by other procedures. We could find a case where excellent feedback is being provided but the procedures called for by the training program are not followed. That is, the function is being filled, but the procedures are not followed.

Sufficiency evaluation is important because it gets us away from the assumption that any change in the program is bad. If the users

change the program to bring it more in line with their way of doing things, then the users may have substituted procedures of their own which fill the same function as the procedures invented by the developer. If this were the case, then these procedures or components would be considered implemented and provide input to the trial evaluation. If the alterations are not sufficient then an effort must be made to change the way the users are doing things. This provides input to evaluation of strategies and tactics.

Re-analysis of Fit. Not shown in Figure 3 is the input the various measures of degree-of-implementation provide to the re-analysis-of-fit. The re-analysis identifies implementation problems which are still unresolved and permits the implementation effort to focus on finding strategies and developing tactics to resolve them. Strategies and tactics which focus on already implemented components can be dropped from the implementation effort.

Trial Evaluation

The trial evaluation fills the gap between formative and summative evaluation. A double meaning is intended. It is an evaluation of program effectiveness during the trial period of implementation (Roberts-Gray & T. Gray, 1983) and it is a trial summative evaluation. The trial evaluation is expected to be constantly revised or redone during the implementation period. Because of this it is not evaluating a fixed target, that is, the program itself changes (the hope is that more becomes implemented) during the course of the evaluation. The interpretation of trial evaluation results must be weighted by

considering what program components are implemented currently. The more "must have" components that are implemented, the more the trial evaluation data can be interpreted as a true summative evaluation.

In bridging the gap between formative and summative evaluation, the trial evaluation has characteristics of both these evaluations as well as serving as part of the implementation effort. In some ways the trial evaluation can be regarded as an *in vivo* formative evaluation. Just as the formative evaluation provides feedback to the developers as to what works under laboratory conditions, the trial evaluation can provide feedback needed to *finetune* the innovation for maximum effectiveness under real-world conditions. Additionally, the best time to make changes in the program is during implementation while users are still receptive to change.

Where formative evaluation provides an assessment of program effectiveness under ideal conditions, trial evaluation assesses the effectiveness of the program-as-implemented-currently. The gap between current and expected effectiveness can act as an argument for continued attention to the implementation effort. While the same gap might be found in comparing a summative with a formative evaluation, by the time the summative evaluation is conducted most implementation efforts have stopped. Intuitively it seems much easier to keep the implementation effort going than to restart it once stopped.

Finally, in many cases the summative evaluation can be built upon the design, methodology, and instruments used in the trial evaluation. If techniques such as time-series analyses are used, then even the data

collected in the trial evaluation may be included in the summative evaluation.

Strategies and Tactics

An early reading on how well a program is being implemented can be gained from evaluating the achievement of strategic goals. If the theory-driven selection of strategies pinpointed certain strategic goals as necessary to resolve certain PIPs, then achievement of these goals logically precede full use of the innovation. A close monitoring of strategic goals may allow the monitor to recommend changes in implementation tactics before the organization's inital burst of enthusiasm (and money allocated for implementation) is spent.

Insert Figure 4 about here

If the initial evaluation of degree-of-implementation has been completed, then the only strategic goals evaluated are those pertaining to components not yet implemented. In the example given earlier, if evaluation of degree-of-implementation reveals that controllers are not interrupting tactical field exercises to give feedback to trainees, then the PIP identified above has been resolved, and evaluation of the strategies proposed to resolve the PIP would not be conducted.

In contrast, if a component is not implemented but the relevant strategic goals are achieved, then additional strategies must be selected and tactics choosen. (See Figure 4.) However, if a component is not implemented and the strategic goals have not been acheived, then

(see Figure 4) an evaluation of tactics is conducted.

In evaluating tactics the basic question is whether the tactic was well executed or not. For example, to educate controllers on the virtues of unobtrusive behavior a training manual might be produced (a tactic). The monitors can ask whether the manual actually presents arguments in favor of unobtrusive behavior, whether the reading level and format is appropriate for the intended audience, whether controllers actually received the manuals, and whether the controllers have read the manuals.

If the tactics were well executed, but the strategic goal was not acheived, then there is a need to develop and execute a new set of tactics to achieve that strategic goal. In contrast, if the tactics are NOT well executed, then revision or re-execution is required.

The evaluation of tactics may be empirical or analytic. For example, in writing training pamphlets, the "nice to know" information is often confused with the "must know" information. In this example, the monitor team may first have to determine what is "must have" and what is "nice to know." Then they would perform a critical reading of the pamphlet to determine if the "must know" information is adequately represented. For other tactics the monitor may have to perform a mini-program evaluation. For example, for new Army training programs, it is common to send a mobile team to each post to train-the-trainers. In this case the monitor might want to assess whether course graduates can indeed train others.

Implementation monitoring is an iterative process. By the time the

evaluation of strategies and tactics is completed, an update of the degree-of-implementation evaluation or a re-analysis-of-fit will have been started. Either action will restart the cycle and eventually lead to a re-evaluation of strategies and tactics.

Utility of Framework

The validation of the framework for implementation monitoring lies in its utility. I see three related uses for this framework.

First, the framework may be used to guide monitoring efforts. At present, those anticipating problems in implementing an innovation are faced with a grab bag of rules-of-thumb and warnings. Little systematic guidance is provided. The present framework should be of use to all workers in the field.

Second, the framework may be used in a retrospective analysis of implementation problems. In this way it provides a common basis to organize implementation studies and facilitate comparison across studies.

Third, the framework may be used to compare and contrast theories of implementation. This can occur in either the retrospective or monitoring modes. The factors suggested by different theories can be fed into the analysis-of-fit and selection of strategies and tactics The data gathered can be used to determine which theories were more useful and suggest factors omitted by current theories.

Summary

Implementation is not an event. It is a process. To ensure that a new program is used fully a process of implementation planning and monitoring must be executed.

The framework discussed in this paper provides a systematic procedure for ensuring that implementation planning is complete and is well executed. The framework defines certain types of background knowledge as important prerequisites to implementation monitoring: practical knowledge of the innovation's procedures and resource needs; practical knowledge of the implementing organization; knowledge of relevant domain theories; and knowledge of implementation theories. This mixture of practical and theoretical knowledge is combined to yield three rational analyses: a ranking of program components, an analysis-of-fit between the innovation and organization, and a characterization of the implementation effort in terms of the strategies and tactics required to resolve areas of misfit (PIPs). While potentially useful in themselves, these rational analyses define the questions asked in the three empirical evaluations.

Evaluation of the degree-of-implementation looks at the fidelity of the implementation to developer's ideal and assesses the sufficiency of any adaptations. The trial evaluation assesses how well the innovation is working at its current level of implementation. It provides an impetus to continued implementation efforts as well as feeding into the summative evaluation when implementation efforts cease. Evaluation of strategies and tactics provide both an early reading on the

effectiveness of the implementation effort and the feedback needed to increase the degree-of-implementation.

To conclude, the framework for implementation monitoring defines a process for implementing new programs in organizations. An optimal use of the framework would be to improve the use of new programs by implementation monitoring. A sufficient use of the framework is to raise the awareness in organizations (and among evaluators) of the issues involved in implementing new programs.

Footnotes

[1]Some of the ideas discussed in this paper were presented in rudimentary form during presentations by the author to American Psychological Association and Evaluation Research Society/Evaluation Network (W. Gray, 1982; W. Gray & Roberts-Gray, 1982).

References

Berman, P. (1978). The study of macro- and micro-implementation. Public Policy, 26, 157-184.

Berman, P. (1980) Thinking about programmed and adaptive implementation: Matching strategies to situations. In H. Ingram & D. Mann (Eds.), Why policies succeed or fail (pp.205-230). Beverly Hills, CA: Sage.

Chen, H., & Rossi, P. H. (1983) Evaluating with sense: The theory driven approach. Evaluation Review, 7, 283-302.

Emshoff, J. G., Mayer, J., Gottschalk, R., Blakely, C., & Roitman, D. (1984, October). Innovation in public sector organizations: Measuring fidelity of implementation and program effectiveness. In W.D. Gray (Chair), Implementation: A key to high organizational performance. Panel conducted at the meeting of the Evaluation Research Society, San Francisco, CA.

Fisher, C. W., Berliner, D. C., Filby, N. N., Marliave, R., Cahen, O.S., & Dishaw, M. M. (1981). Teaching behaviors, academic learning time, and student achievement: An overview. Journal of Classroom Interaction, 17, 2-15.

Fullan, M., & Pomfret, A. (1977). Research on curriculum and instruction implementation. Review of Educational Research, 47, 335-397.

Gray, T. (1981). Implementing innovations: A systems approach to integrating what is known. Journal of Technology Transfer, 6, 19-32.

Gray, W. D. (1983). Engagement simulation: A method of tactical team training. Training & Development Journal, 37(7), 29-34.

Gray, W. D. (1982, August). Life cycle evaluation: A "cradle-to-maturity" approach for Army training programs. Paper presented at the meeting of the American Psychological Association, Washington, DC.

Gray, W. D., & Roberts-Gray, C. (1982, October). Implementing Army training programs: Translating model into action. Paper presented at the meeting of the Evaluation Research Society, Baltimore, MD. (ERIC Document Reproduction Service No. ED 230 591)

Hall, G.E., & Loucks, S.F. (1977). A developmental model for determining whether the treatment is actually implemented. American Educational Research Journal, 14, 263-276.

Leinhardt, G. (1980). Modeling and measuring educational treatment in evaluation. Review of Educational Research, 50, 393-420.

Roberts-Gray, C. (1983). On Closing the Implementation Gap: Symposium Proceedings (Res. Rep., in press). Alexandria, VA: Army Research Institute.

Roberts-Gray, C., & Gray, T. (1983). Implementing innovations: A model to bridge the gap between diffusion and utilization. Knowledge: Creation, Diffusion, Utilization, 5, 213-232.

Scheirer, M. A. (1981). Program implementation: The organizational context. Beverly Hills: Sage Publications.

Scheirer, M. A., & Rezmovic, E. L. (1983). Measuring the degree of program implementation: A methodological review. Evaluation Review, 7.

Scott, T. (1983). Implementing innovations in the Army: A case study. In C. Roberts-Gray (Ed.), On closing the implementation gap: Symposium proceedings (pp.9-17). (Research Report, in press). Alexandria, VA: Army Research Institute.

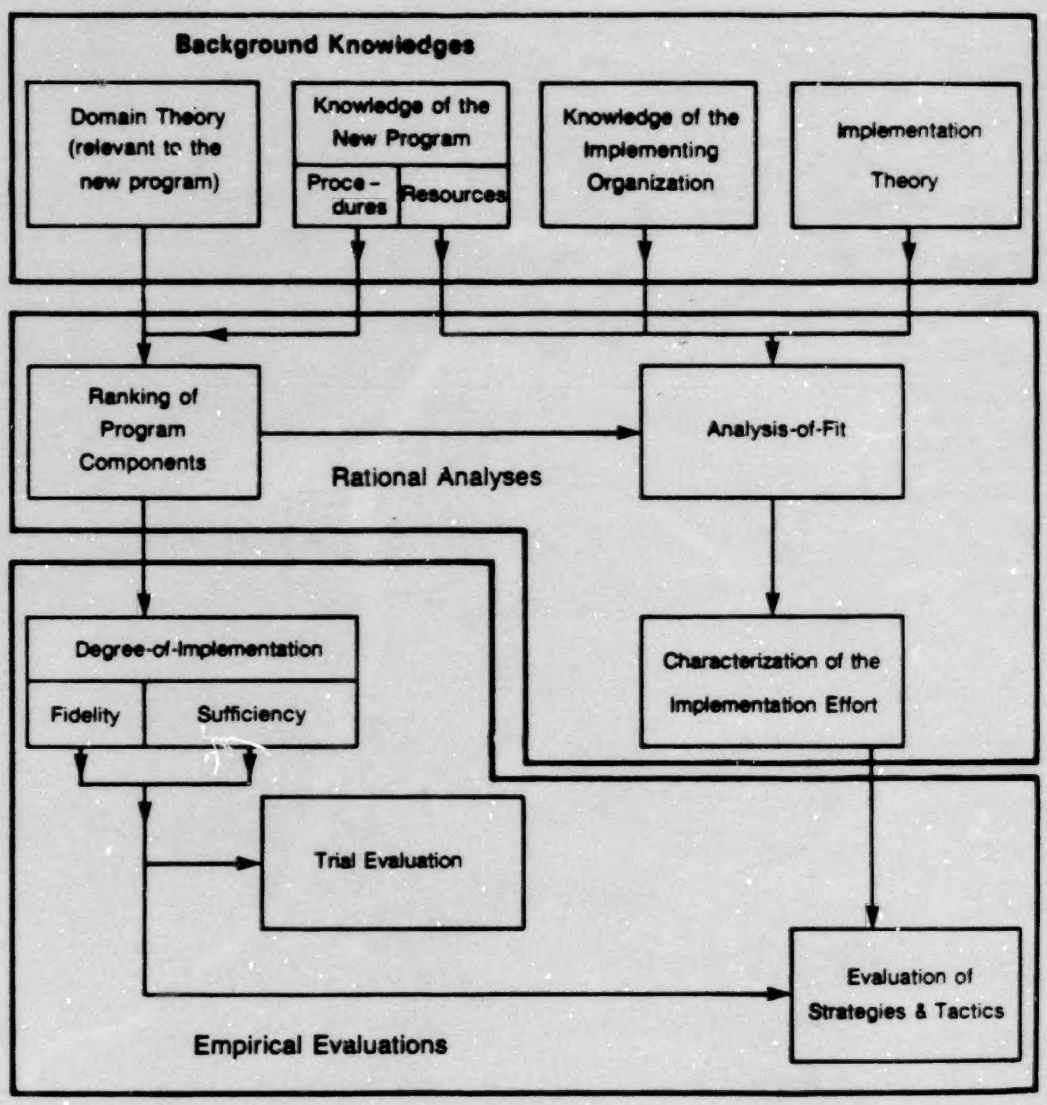
Figure Caption

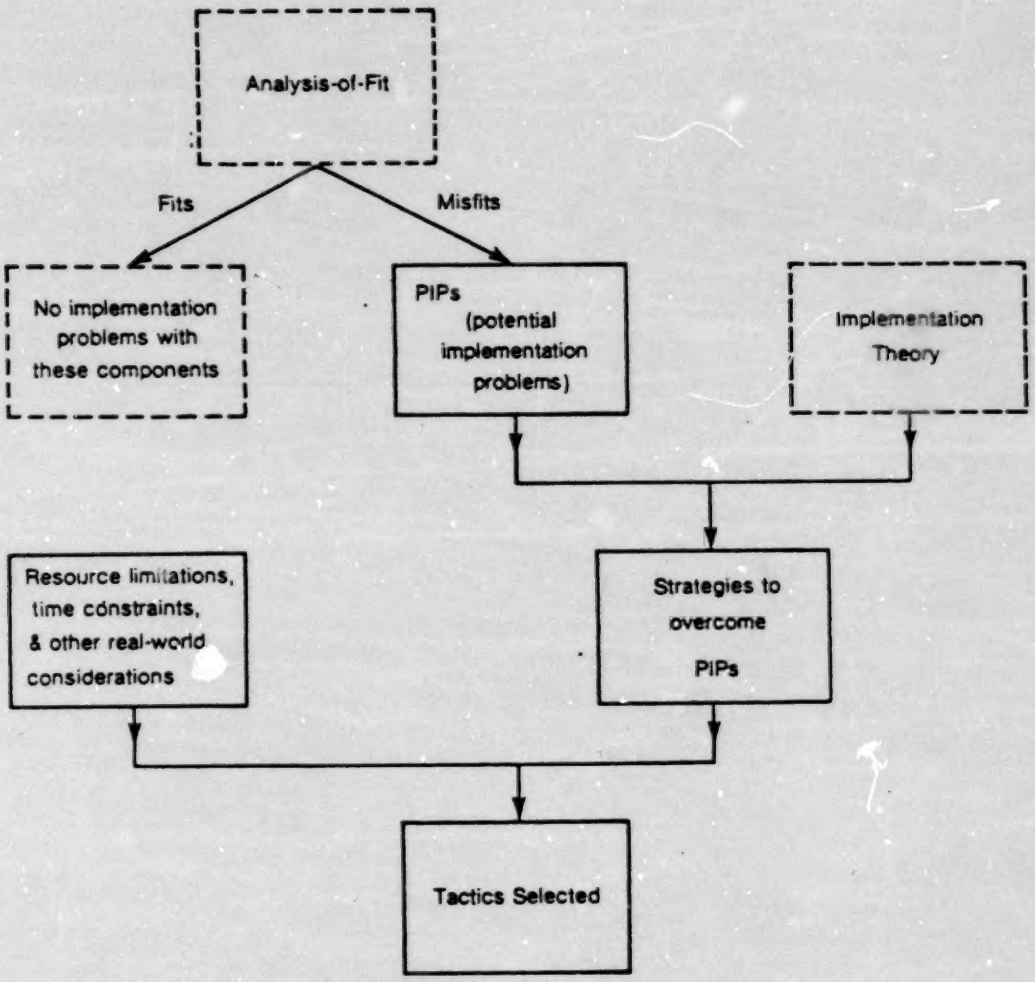
Figure 1. Framework for Implementation Monitoring

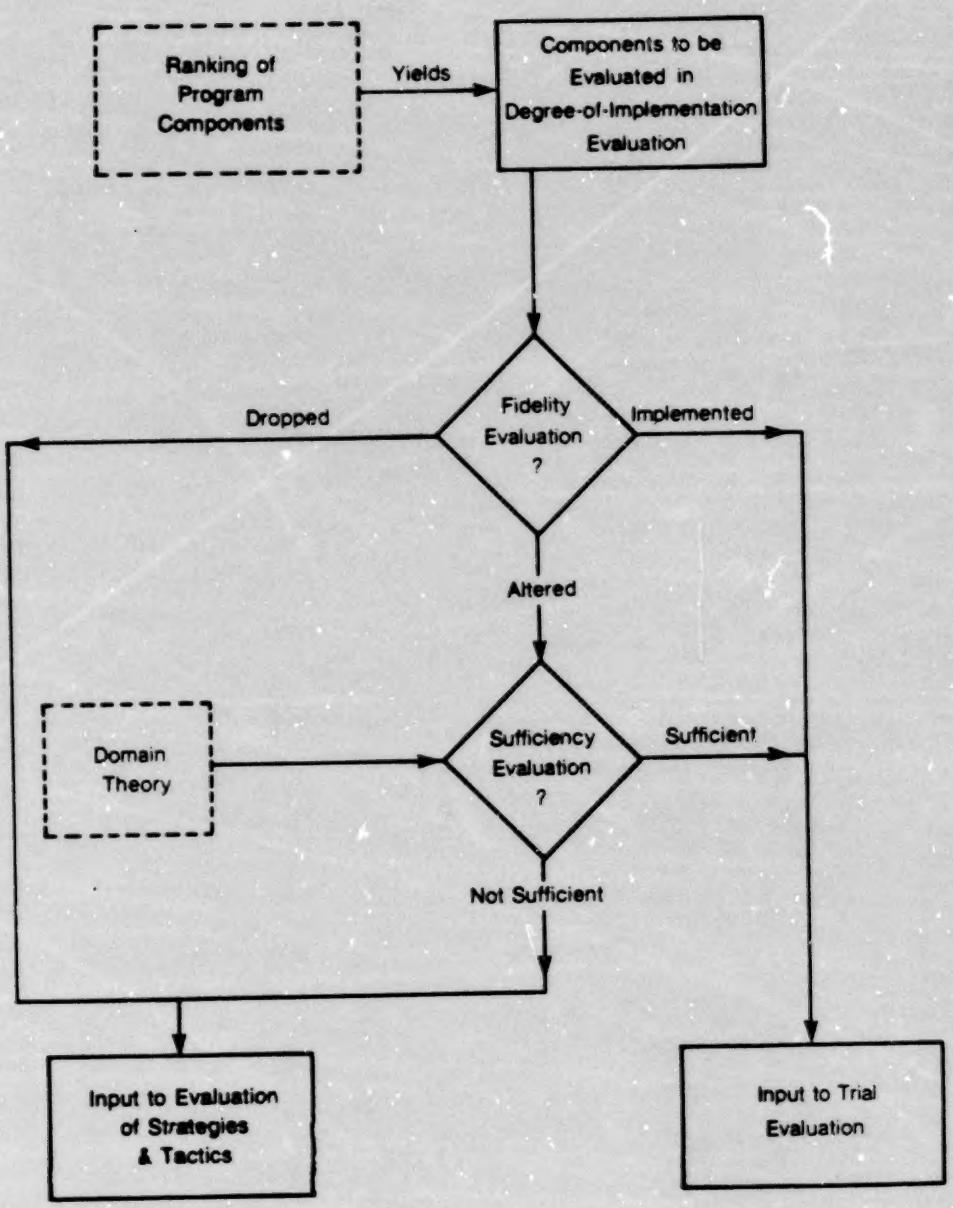
Figure 2. Characterization of the Implementation Effort

Figure 3. Degree-of-Implementation

Figure 4. Evaluation of Strategies & Tactics


Background Knowledges
Domain Theory (relevant to the new program)
Knowledge of the New Program
Proce-dures
Resources
Knowledge of the Implementing Organization
Implementation Theory
Ranking of Program Components
Rational Analyses
Analysis-of-Fit
Degree-of-Implementation
Fidelity
Sufficiency
Characterization of the Implementation Effort
Trial Evaluation
Evaluation of Strategies & Tactics
Empirical Evaluations


Analysis-of-Fit
Fits
Misfits
No implementation problems with these components
PIPs (potential implementation problems)
Implementation Theory
Resource limitations, time constraints, & other real-world considerations
Strategies to overcome PIPs
Tactics Selected


Ranking of Program Components
Yields
Components to be Evaluated in Degree-of-Implementation Evaluation
Fidelity Evaluation ?
Dropped
Implemented
Altered
Domain Theory
Sufficiency Evaluation ?
Sufficient
Not Sufficient
Input to Evaluation of Strategies & Tactics
Input to Trial Evaluation

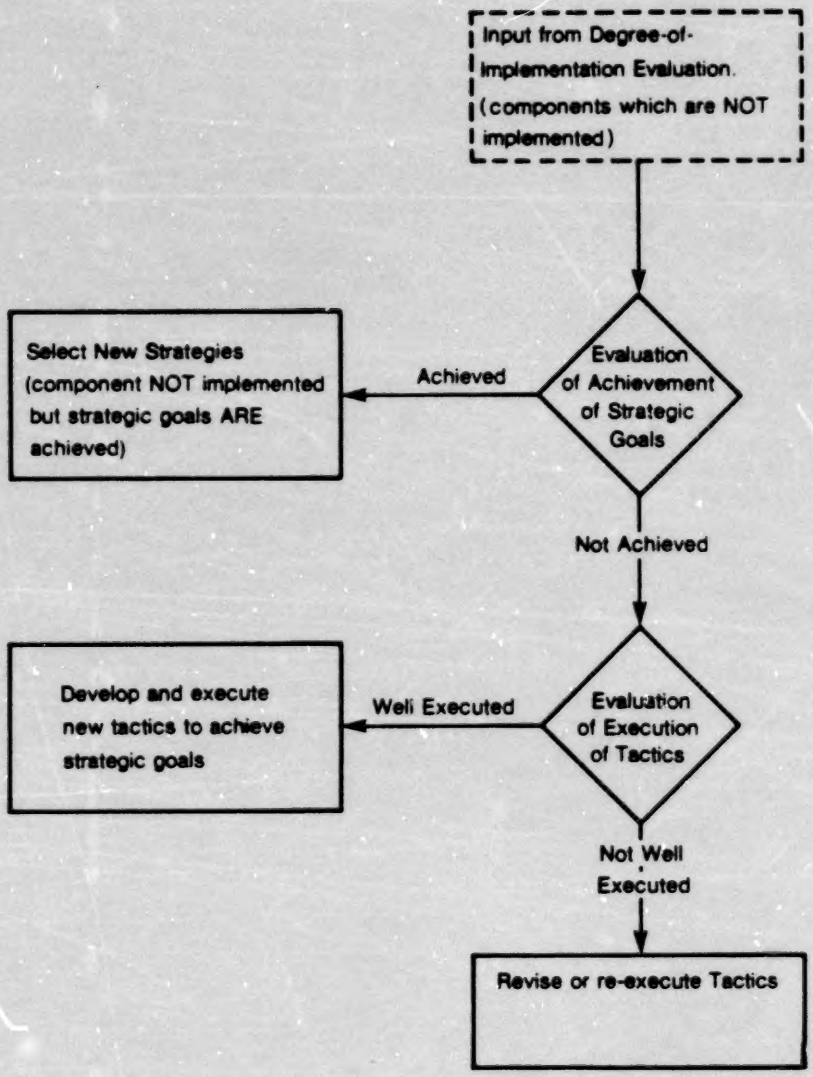

Input from Degree-of-Implementation Evaluation. (components which are NOT implemented)
Evaluation of Achievement of Strategic Goals
Achieved
Select New Strategies (component NOT implemented but strategic goals ARE achieved)
Not Achieved
Evaluation of Execution of Tactics
Well Executed
Develop and execute new tactics to achieve strategic goals
Not Well Executed
Revise or re-execute Tactics

Table 1

Framework Combining Army Guidance on Field Training With Model of Teacher Behavior

---

PREPARATION

- Trainers
- Soldiers (trainees)
- Equipment
- Training Area

PRESENTATION

- Planned Explanations
  - Training Objectives of Exercise
  - "Operations Order"
- Unplanned Explanations
  - Responsiveness to trainee questions
- Structure of Exercise
  - Scenario structured to support training objectives
  - Area Selected, positions picked, force ratios, etc. in support of training objectives

PRACTICE

- Monitoring
  - Observation of trainee activities
    - -- enforcement of rules of engagement (classroom control)
    - -- observation and recording of trainee tactical performance
  - Questioning students
- Feedback
  - Concerning tactical proficiency
  - To control attention to task

DIAGNOSIS & PRESCRIPTION

- Level at Which Diagnosis and Prescription Takes Place (i.e., echelon which manages instruction)
- Role, If Any, Trainers Have in Selecting or Changing Tasks Once in the Field
- Appropriateness of Instruction
  - To trainee needs
  - To trainee skill level

Table 2

Theory-based Component Analysis: Feedback procedures for an Army tactical training program

III Practice

B. FEEDBACK

1. Concerning tactical proficiency

a. malfunctioning equipment: replaced or repaired during exercise (Note: this is to ensure feedback from the laser devices which simulate weapons effects)

b. during exercise: no feedback provided to individuals or groups concerning tactical proficiency (Note: controller stays as unobtrusive as possible to encourage realism)

c. after exercise--After Action Review conducted (Note: "socratic" procedure meant to encourage participation of all trainees)

(1) Feedback on collective performance is emphasized

(2) Feedback on individual performance minimized

(3) Frequency of After Action Review and timing (for example, immediately after each exercise, once a day, and so on)

2. To control trainee attention to task

a. on-the-spot corrections to enforce rules-of-engagement

b. control of the After Action Review process (to focus discussion on relevenat training objectives, and minimize "who-shot-whom" discussions)